DIGITALCOMMONS @WAYNESTATE

Journal of Modern Applied Statistical Methods

Volume 5 | Issue 1 Article 12

5-1-2006

Penalized Splines For Longitudinal Data With An Application In AIDS Studies

Hua Liang
University of Rochester Medical Center, hliang@bst.rochester.edu

Yuanhui Xiao
Georgia State University, matyxx@langate.gsu.edu

Journal of Modern Applied Statistical Methods
May, 2006, Vol. 5, No. 1, 130-139

1538 – 9472/06/$95.00

Penalized Splines For Longitudinal Data With An Application In AIDS Studies

Hua Liang
Department of Biostatistics and Computational Biology
University of Rochester Medical Center

Yuanhui Xiao
Department of Mathematics and Statistics
Georgia State University

A penalized spline approximation is proposed in considering nonparametric regression for longitudinal data. Standard linear mixed-effects modeling can be applied for the estimation. It is relatively simple, efficiently computed, and robust to the smooth parameters selection, which are often encountered when local polynomial and smoothing spline techniques are used to analyze longitudinal data set. The method is extended to time-varying coefficient mixed-effects models. The proposed methods are applied to data from an AIDS clinical study. Biological interpretations and clinical implications are discussed. Simulation studies are done to illustrate the proposed methods.

Key words: Repeated measurements, varying-coefficient models, AIDS, ACTG315

Introduction

Recently, nonparametric regression has been used to analyze longitudinal data, which arise frequently in clinical trials and biological research and cannot be analyzed by traditional parametric approaches. The aims of nonparametric regression analysis include exploration of curves for a particular population and individual characteristic by introducing a mixed-effects framework. For parametric longitudinal data, for surveys, see Diggle, Liang and Zeger (1994), Davidian and Gilti-nan (1995), Vonesh and Chinchilli (1996) among others. Mixed-effects models provide a useful and flexible framework in which population characteristics are modeled as fixed effects, while individual variation is modeled as a random effect. Parametric mixed-effects models such as linear mixed-effects (LME) models (Laird & Ware 1982, Ware 1985, Diggle, et al. 1994) and nonlinear mixed-effects models (Davidian & Giltinan 1995, Vonesh & Chinchilli 1996) are widely used in longitudinal data analysis. Shi, Weiss, and Taylor (1996) and Rice and Wu (2001) proposed a nonparametric mixed-effects model for longitudinal data:

$$y_i(t) = \eta(t) + v_i(t) + \varepsilon_i(t), \quad i = 1,2, ..., n, \tag{1}$$

where $\eta(t)$ models the population mean function, also called the fixed-effect or population curve; $v_i(t)$ models individual variations from $\eta(t)$ (these variation are called random-effect curves); $\varepsilon_i(t)$ are measurement errors; and $y_i(t)$ are response processes. The $v_i(t)$ and $\varepsilon_i(t)$ are assumed to be independent. $v_i(t)$'s can be considered as realizations of a zero mean process with a covariance function $\gamma(s,t) = E\{v_i(s)v_i(t)\}$, and $\varepsilon_i(t)$ can be regarded as realizations of an uncorrelated zero mean process with a variance function $\sigma^2(t)$. Let t_{ij}, $j = 1,2, ..., m_i$, be the design time points for the *i-th* individual, then model (1) becomes

Hua Liang is Associate Professor. His research interests are in methodologies for analyzing data in biomedical research, including longitudinal studies and clinical trials. Email: hliang@bst.rochester.edu. Yuanhui Xiao is Assistant Professor. His research interests are in time series analysis, statistical computation, statistical software development. E-mail: matyxx@langate.gsu.edu.

$$y_i(t_{ij}) = \eta(t_{ij}) + v_i(t_{ij}) + \varepsilon_i(t_{ij}), \quad j = 1,2,\ldots,m_i; \quad i = 1,2,\ldots,n, \tag{2}$$

where n is the number of subjects and mi is the number of measurements from subject i. For convenience, y_{ij} is denoted as being equal to $y_i(t_{ij})$ and ε_{ij} as being equal to $\varepsilon_i(t_{ij})$.

The primary goal is to estimate the fixed-effect (population) curve $\eta(t)$ and random-effect curves $v_i(t)$ or individual curves $s_i(t) = \eta(t) + v_i(t)$, for $i = 1,2, \ldots, n$. The mean function $\eta(t)$ is important because it reflects the overall trend or progress of an underlying population process and can be used as an important index for the population response to a drug or a treatment in a clinical or biomedical study. The estimation of $v_i(t)$ or $s_i(t)$ is also important. The estimates of $v_i(t)$ are crucial for the estimation of the covariance of $y_i(t)$, which, in turn, can be used to better the estimate of the population curve $\eta(t)$ (see later sections). Because an individual curve $s_i(t)$ may represent an individual response to a treatment in a study, a good estimate of $s_i(t)$ may help investigators to make a better decision about individual treatment. The estimates of individual curves $s_i(t)$ are also useful if the investigators wish to group or classify the subjects on the basis of individual response curves.

Several methods have been proposed for the nonparametric modeling of longitudinal data. Diggle and Hutchison (1989), Altman (1990), Hart (1991), Rice and Silverman (1991) and others proposed modifications to criteria for selection of smoothing parameters. These modifications include leave-one-subject-out cross-validation (CV) or generalized cross-validation (GCV) to indirectly account for the correlations among data. Zhang et al. (1998) considered the correlation structure of longitudinal data in their smoothing spline semi-parametric mixed-effects models, but only the population curve (mean function) is modeled non-parametrically.

Wang (1998a, b) included the correlation in a mixed-effects smoothing spline models, but the special correlation structure of longitudinal data was not emphasized. Hart and Wehrly (1986) and Fan and Zhang (2000) suggested a two-step approach (local averaging or local regression first, then smoothing) to indirectly account for the data correlation. Hoover et al. (1998) and Wu, Chiang and Hoover (1998) proposed a standard local polynomial kernel method for varying-coefficient model with longitudinal data. Lin and Carroll (2000) propose a local polynomial generalized estimating equation (GEE) method for clustered data that may also be used to estimate the population curve $\eta(t)$ in our model. More recently, Wu and Zhang (2002) suggested that $\eta(t)$ and $v_i(t)$ be estimated simultaneously by combining LME models and local polynomial techniques, and they propose new bandwidth selection methods that are hybrid approaches of leave-one-subject-out and leave-one-point-out CV.

Although all of these approaches have demonstrated promise, several potential weaknesses exist.

(a) All these existing methods, except that of Wu and Zhang (2002), did not consider estimating the random-effect curves $v_i(t)$ or individual curves $s_i(t)$, which are very important in the application of the models to data from clinical and biological studies.

(b) The approach of Wu and Zhang (2002) has been shown to be more efficient than the other approaches, and the authors considered individual curves $s_i(t)$, but the computation of their methods is very expensive and sometimes unstable for bandwidth variation.

(c) Even if these weaknesses are ignored, the selection of smoothing parameters depends heavily upon selection criterion such as AIC, BIC or GCV.

Here, a new method is proposed to simultaneously estimate $\eta(t)$ and $v_i(t)$ by combining LME models (Laird & Ware 1982) and penalized techniques (Carroll & Ruppert 1999). The resulting estimators are called penalized spline LME (PSLME) estimators. This approach overcomes the above weakness, and is simple, easily and quickly implemented and robust to smoothing parameters.

An approach similar to the one proposed here has been used for common nonparametric regression. Parise, Ruppert, Ryan and Wand (2001) proposed penalized spline model to study the relationship between animal body weight and tumor onset by incorporating variation from one experiment to another. A similar mixed model was used to analyze the data from a study of the Utah Valley respiratory health/air pollution study by Coull, Schwartz and Wand (2001), and from a study of ragweed pollen data by Coull, Ruppert and Wand (2001).

The rest of the paper is organized as follows. Section 2 shows the derivation of the PSLME estimators and an extension to time varying coefficient mixed-effects model. As an illustration, an application of the model to a data set from an AIDS study is shown in section 3.1. A simulation study is presented in section 3.2. Some discussions are given in section 4.

Estimation Framework

Before the estimation framework is established, the principle of penalized spline for classic non-parametric regression is briefly introduced. More details were described by Ruppert and Carroll (1999).

The penalized least-squares estimator

The data (X_i, Y_i) follow $Y_i = m(X_i) + e_i$ for i = 1, 2, …, n, where X_i is univariate. To estimate m, β is equal to $(\beta_0, \beta_1, \ldots, \beta_p)^T$ and a regression spline model

$$m(x;\beta) = \beta_0 + \beta_1 x + \cdots + \beta_p x^p + \sum_{k=1}^{K} b_k (x - \zeta_k)_+^p$$

is used to approximate $m(x)$, where $p \geq 1$ is an integer and $\zeta_1 < \cdots < \zeta_K$ are fixed knots, $a_+ = \max(a, 0)$. The traditional method of "smoothing" the estimate is knot selection. The estimator $\hat{\beta}(\alpha)$ of β is defined as the minimizer of

$$\sum_{i=1}^{n} \{Y_i - m(X_i; \beta)\}^2 + \alpha \sum_{k=1}^{K} b_k^2 , \tag{3}$$

where α is a smoothing parameter.

As shown by Brumback, Ruppert and Wand (1999), the estimator $\hat{\beta}(\alpha)$ based on equation (3) is equivalent to the estimator of β based on an LME model

$$y = X\beta + Zb + \varepsilon ,$$

where

$$X = \begin{pmatrix} 1 & X_1 & \ldots & X_1^p \\ 1 & X_2 & \ldots & X_2^p \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_n & \ldots & X_n^p \end{pmatrix},$$

$$Z = \begin{pmatrix} (X_1 - \zeta_1)_+^p & (X_1 - \zeta_2)_+^p & \ldots & (X_1 - \zeta_K)_+^p \\ (X_2 - \zeta_1)_+^p & (X_2 - \zeta_2)_+^p & \ldots & (X_2 - \zeta_K)_+^p \\ \vdots & \vdots & \ddots & \vdots \\ (X_n - \zeta_1)_+^p & (X_n - \zeta_2)_+^p & \ldots & (X_n - \zeta_K)_+^p \end{pmatrix}$$

$$b = (b_1, \ldots, b_K)^T \sim N(0, \sigma_b^2),$$

$$\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)^T \sim N(0, \sigma_\varepsilon^2),$$

and

$$\alpha = \sigma_\varepsilon^2 / \sigma_b^2 .$$

This fact implies that penalized spline smoother under the framework in equation (3) is equivalent to a standard LME. The solution can be obtained through the use of an LME macro available for S-PLUS software. The penalized

parameter α is automatically estimated as $\hat{\alpha} = \hat{\alpha}_\varepsilon^2 / \hat{\alpha}_b^2$ by a restricted maximum likelihood (RML) approach.

Estimation Procedures for Model (3)

Motivated by the idea stated in Section 2.1, an estimation approach is proposed as follows. First, $\{(t_{ij}, Y_{ij})\}$ ($j = 1,2, ..., m_i$ and $i = 1,2, ..., n$) are the data drawn from the model in (2). The fixed effects functions $\eta(t)$ are approximated by

$$\tilde{\eta}(t, \beta, u) = \sum_{k=0}^{p} \beta_k t^k + \sum_{k=1}^{K} u_k (t - \zeta_k)_+^p$$

and those of $v_i(t)$ are approximated by

$$\tilde{v}_i(t, b_i, w_i) = \sum_{k=0}^{p} b_{ik} t^k + \sum_{k=1}^{K} w_{ik} (t - \zeta_k)_+^p$$

Here

$$\beta = (\beta_0, \ldots, \beta_p)^T, \quad u = (u_1, \ldots, u_K)^T,$$
$$b_i = (b_{i0}, \ldots, b_{ip})^T, \quad w_i = (w_{i1}, \ldots, w_{iK})^T.$$

Assume that $\{u_k\} \sim N(0, \sigma_u^2), \{b_{ik}\} \sim N(0, \sigma_{b,k}^2)$ and $\{w_{ik}\} \sim N(0, \sigma_w^2)$ for $k = 1, \ldots, K$ and $i = 1, ..., n$. Then $\tilde{\eta}(t, \beta, u) + \tilde{v}_i(t, b_i, w_i)$ is the individual curve of the i^{th} subject. Define the following matrix notation.

$$X_i = \begin{pmatrix} 1 & t_{i1} & \ldots & t_{i1}^p \\ 1 & t_{i2} & \ldots & t_{i2}^p \\ \vdots & \vdots & \ddots & \vdots \\ 1 & t_{im_i} & \ldots & t_{im_i}^p \end{pmatrix},$$

$$Z_i = \begin{pmatrix} (t_{i1} - \zeta_1)_+^p & (t_{i1} - \zeta_2)_+^p & \ldots & (t_{i1} - \zeta_K)_+^p \\ (t_{i2} - \zeta_1)_+^p & (t_{i2} - \zeta_2)_+^p & \ldots & (t_{i2} - \zeta_K)_+^p \\ \vdots & \vdots & \ddots & \vdots \\ (t_{im_i} - \zeta_1)_+^p & (t_{im_i} - \zeta_2)_+^p & \ldots & (t_{im_i} - \zeta_K)_+^p \end{pmatrix},$$

$$\eta_i(t_i) = \{\eta_i(t_{i1}), \ldots, \eta_i(t_{im_i})\}^T,$$
$$y_i = (y_{i1}, \ldots, y_{im_i})^T, \quad X = (X_1^T, \ldots, X_n^T)^T,$$
$$y = (y_1^T, \ldots, y_n^T)^T,$$
$$\Lambda = diag(X_1, \ldots, X_n), \quad Z = (Z_1^T, \ldots, Z_n^T)^T,$$
$$\Gamma = diag(Z_1, \ldots, Z_n),$$
$$b = (b_1^T, \ldots, b_n^T)^T, \quad w = (w_1^T, \ldots, w_n^T)^T.$$

The approximation of the model in (2) can be rewritten as

$$y = X\beta + \Lambda b + Zu + \Gamma w + \varepsilon$$

This standard LME has unknown population parameters β and unknown individual effects b, u and w. The estimates $\hat{\beta}$, $\hat{b}$, $\hat{u}$ and $\hat{w}$ of the parameter vector can be easily given closed forms, and the well-developed SAS and S-plus macros can be directly applied for computation. As a consequence, population and individual curves can be obtained from the estimates $\tilde{\eta}(t, \hat{\beta}, \hat{u})$ and $\tilde{v}_i(t, \hat{b}_i, \hat{w}_i)$.

For a common penalized spline, the penalty parameter α and the number of knots K must be selected. Relatively speaking smoothing is controlled by the penalty parameter α, and the number of knots K is not a crucial parameter. See also Ruppert (2002) for a detailed discussion. As indicated in section 2.2, the formulation of mixed-effects model automatically derives an estimated of α. Only K needs to be specified. Computation experience indicates $\max(10, n/4)$ is a good choice as a value of K and that the results are very insensitive to different values of K. The knots are then at equally spaced sample quantiles of $\{t_{ij}\}$.

Extension to Time Varying-coefficient Models

As an effective approach to reduce curse of dimensionality suffered in high-dimension non-parametric regression, time varying-coefficient models were first proposed in longitudinal data structure by Hoover, Rice, Wu

and Yang (1998) and Wu, Chiang and Hoover (1998). The standard time-varying coefficient models (Hoover et al. 1998, Wu et al. 1998) can be written as

$$y_i(t) = c^T(t) + \eta_i(t) + \varepsilon_i(t), \quad i = 1, \ldots, n, \tag{4}$$

Where $c(t) = \{1, c_1(t), \ldots, c_L(t)\}^T$ and $\eta_i(t) = \eta(t) + v_i(t)$ with $\eta(t) = \{\eta_0(t), .., \eta_L(t)\}^T$ and $v_i(t) = \{v_{0i}(t), \ldots, v_{Li}(t)\}^T$. The functions $\eta_l(t)$ and $\eta_l(t) + v_{li}(t)$ indicate the population and individual effects of $c_l(t)$ for subject i.

Both smoothing spline and local polynomial kernel regression methods are proposed by Hoover et al. (1998). Alternatively, Fan and Zhang (2000) proposed a two-step method for the same model. However, none of these methods efficiently considered the important features of longitudinal data such as between-subject and within-subject variation, and the special correlation structure of longitudinal data. Lin and Carroll (2000), however, showed that specifying the correlation structure when using kernel methods to estimate the nonparametric function results an asymptotically less efficient estimator than the one obtained assuming independence among repeated measures. Welsh, Lin and Carroll (2000) showed regression and smoothing splines do not suffer from this difficulty.

Local polynomial estimates of Hoover et al. (1998) rely upon one bandwidth to smooth all coefficient curves, but these estimates may not be enough to capture smoothness of all coefficient curves simultaneously. The smoothing spline method of Hoover et al. (1998) permits the use of multiple smoothing parameters, but the computation is very intensive even only a single smoothing parameter is included when the number of distinct observation time is large.

More recently, Liang, Wu and Carroll (2003) proposed a global fitting method for a varying-coefficient model based on basis spline approximation. The purpose of their method is to approximate the coefficient functions by the basis spline. Their approach is shown to be simple to estimation and inference for the timing varying coefficient models.

Approximate $\eta_l(t)$ and $v_{li}(t)$ by using the following:

$$\tilde{\eta}_l(t, \beta_l, u_l) = \sum_{k=0}^{p_l} \beta_{lk} t^k + \sum_{k=1}^{K_{l1}} u_{lk} (t - \zeta_{lk})_+^{p_l}$$

and

$$\tilde{v}_{li}(t, b_{li}, w_{li}) = \sum_{k=0}^{q_l} b_{lik} t^k + \sum_{k=1}^{K_{l2}} w_{lik} (t - \zeta_{lk})_+^{q_l}$$

for $l = 1, \ldots, L$. After notation similar to that in section 2.2 is introduced, model (4) can be approximated by

$$y = \sum_{l=1}^{L} X_l \beta_l + \sum_{l=1}^{L} (\Lambda_l b_l + Z_l u_l + \Gamma_l w_l) + \varepsilon.$$

Again, the estimates for all parameters and subsequent population and individual curves can be derived.

Numerical Examples

Analyses of Data from the ACTG 315 Study

ACTG 315 was a single-arm clinical trial in which 53 enrolled subjects with moderately advanced HIV-1 infection received combination antiretroviral therapy consisting of zidovudine, lamivudine, and ritonavir for 48 weeks. The primary objective of the study was to assess whether the treatment was associated with evidence of immunologic restoration. Of the 53 subjects (49 men, 4 women, age range 6-63 years), 44 remained on treatment for at least 9 of the first 12 weeks. Lederman et al. (1998) reported the results of the study after 12 weeks of follow-up. The lower limit of quantification of HIV-1 RNA viral-load is 100 copies/ml. The HIV-1 RNA measures below this limit are not considered reliable; therefore, we censor values that are below 100 copies/ml. HIV-1 RNA measurements were observed on days 0, 2, 7, 10, and weeks 2, 3, 4, 8, 12, 24, and 48 of follow-up.

One aim of the ACTG 315 study is to characterize the viral load trajectory in the population and the individual patients during antiviral treatment. The population estimate of the viral-load trajectory was obtained as a function of

treatment time by using the PSLME method. The estimated curves are presented in Figure 1 in dotted lines. The PSLME method was used to estimate the viral-load trajectories for individual patients. The ability to estimate values for population and individual characteristics is another important advantage of the PSLME method. The individual estimates of viral-load traj ectory for four selected patients are shown in Figure 1, which indicates that individual viral-load trajectories may differ from that estimate for the population. The viral-load trajectory of subject 18 is identical to the viral-load trajectory of the population and the pattern of viral-load trajectory in subject 1 is similar to that of the population, but the difference in magnitude is obvious. Other large differences between individual viral-load trajectory in subjects 23 and 3 5 and that in the population are observed. The estimated trajectories of viral-load in individual patients can provide more accurate information for physicians with which to individualize treatment management for individual patients with AIDS.

To study the relationship between virologic and immunologic responses, repeatedly measured by HIV RNA levels (viral load) and CD4+ cell counts respectively in an AIDS clinical trial ACTG 315, observe that the viral load and CD4+ cell counts are negatively and approximately linearly related in most of the treatment times, but the regression coefficients may not be constant during the whole treatment period. Motivated by this feature of the data, Liang, Wu and Carroll (2003) proposed a mixed-effects varying-coefficient model. The model captures population and individual relationships for the two longitudinal variables. The method proposed above is used to analyze this data set again. In the implementation, set $p = q = 2$, and $K_{l1} = 6$ and $K_{l2} = 10$. Other values were tried, and the results are very stable. The discoveries are similar to what Liang, Wu, and Carroll (2003) obtained. The viral load and CD4+ cell counts are inversely related in the study population during the treatment. However, the strength of the association varies smoothly, where the association is very strong at the beginning of the treatment to the weakest about 4 weeks of treatment. The association gradually recovered and is strongest from week 4 to week 24. See the dotted line in Figure 2 for the population curve.

Figure 2 also shows the individual estimates of $\beta_1(t)$ from four arbitrarily selected patients and the corresponding population estimate of $\beta_1(t)$. Not only the magnitude but also the patterns differ between the population and individual estimates of $\beta_1(t)$ (Figure 2). The pattern for subject 18 is almost identical to that of the population pattern. The patterns for subjects 1 and 47 are similar to the population pattern. However, the viral load and CD4+ cell counts of subject 1 was positive correlated with those of subject 47 during the early treatment stage. For subject 47, there is a negative correlation between viral load and CD4+ cell counts in the later stage. Interestingly we also observe discordance between patterns of the population estimate and individual estimates of $\beta_1(t)$. See pattern for subject 2 shown in Figure 2. Because of the large between-subject variation, the individual estimates become very important in individualizing treatment and care for patients with AIDS.

A Simulation Study

A simulation model is designed as $y_i(t) = \eta(t) + \gamma_i(t) + \varepsilon_i(t)$, where $\eta(t) = 1 + \cos(2\pi t) + \sin(2\pi t)$ and $\gamma_i(t) = a_{i0} + a_{i1}\cos(2\pi t) + a_{i2}\sin(2\pi t)$ with $(a_{i0}, a_{i1}, a_{i2})^T \sim N((0, 0, 0)^T, I_{3\times3})$, and $\varepsilon_i(t) \sim N(0,1)$, for $i = 1, \ldots, n = 20$. The design time points are $t_{ij} = j/(1+m)$ for $j = 1, \ldots, m = 35$. To mimic the unbalanced data feature in longitudinal studies, randomly remove y_{ij} with a probability of r_m = 0.35 (i.e., r_m is the missing rate of the data). Thus, there are an average of 23 observations for each subject and 460 observations in total. Note that the data from different subjects are independent, but the within-subject data are correlated. The within-subject correlation coefficient can be calculated as:

$$\rho_y = corr\{y_i(t), y_i(s)\} = \{1 + \cos 2\pi(t-s)\}/2$$

for $s \neq t$. In this simulation experiment and in later examples, let p = 3 and K = 8 set $\sigma_w^2 = 0$. When a Dell PC machine (2GHz CPU) was used, the computation for the simulation experiment require only 8 seconds. The estimated value of the penalized parameter α is $\hat{\alpha} = 0.034$.

Figure 3 shows the profiles of data for 6 arbitrarily selected subjects. The generated data, the real population curve, the estimated population and individual curves are depicted for comparison. Although the population estimate is similar to the true characteristic of the population, the estimated individual curves more precisely describe individual trends than the estimated population curves. For comparison, this simulation data was set for p = 2 and K = 10,15, 20. The corresponding results are not distinguishable from those in Figure 3.

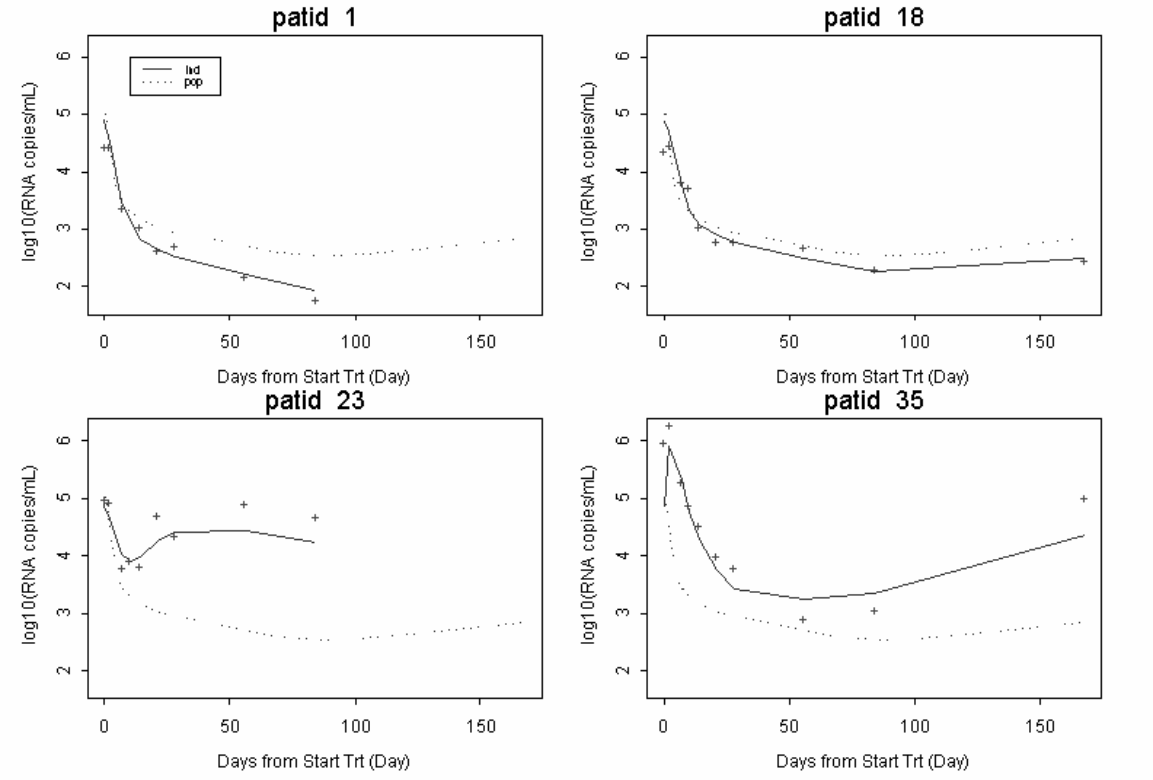

Figure 1.

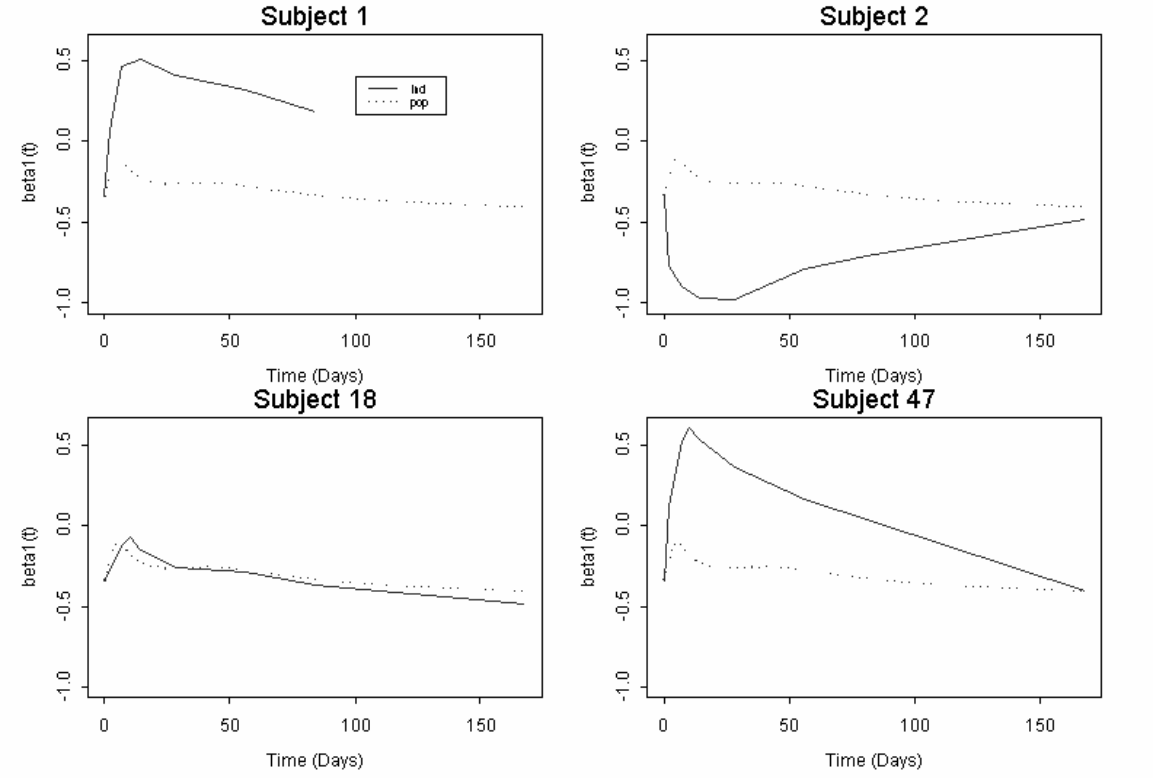
Subject 1
Subject 2
Subject 18
Subject 47
beta1 (t)
Time (Days)
ind
pop
0.5
0.0
-0.5
-1.0
0
50
100
150

Figure 2.

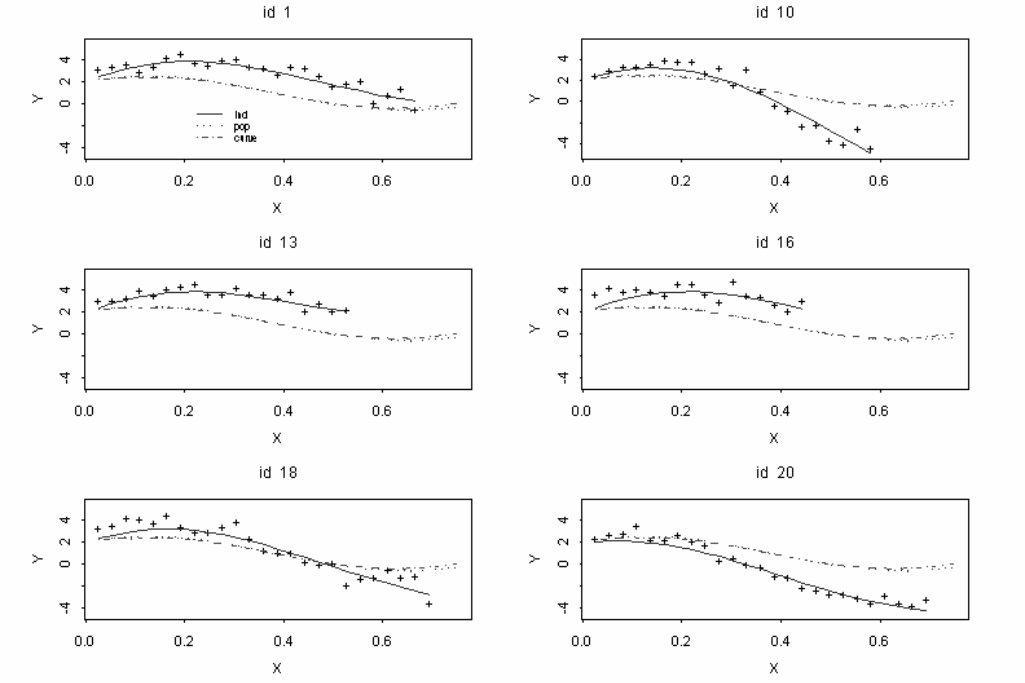

Figure 3.

Conclusion

Considering nonparametric regression modeling for longitudinal data, a very effective routine is proposed by combining a penalized spline technique and LME models. The principal advantage of this approach is that it avoids computational challenges that occur when local kernel smoothing or smoothing spline techniques in which bandwidths or smoothing penalty parameters have to be selected are used. This approach avoids these challenges by using a concern of LME. Penalty parameters were automatically calculated out. Curves for population and individual characteristics are easily derived. The approach is also effective to time varying coefficient mixed-effects models. The method has been shown to be useful in analyzing AIDS data set. It is believed that the approach can be used to other clinical trail or biological data.

References

Altman, N. S. (1990). Kernel smoothing of data with correlated errors. *Journal of the American Statistical Association, 85*, 749-759.

Brumback, B. A. & Rice, J. A. (1998). Smoothing spline models for the analysis of nested and crossed samples of curves (with discussion). *Journal of the American Statistical Association, 93*, 961-994.

Brumback, B. A., Ruppert, D., & Wand, M. (1999). Comments on "Variable Selection and Function Estimation in Additive Nonparametric Regression Using a Data-Based Prior" by Shively, Kohn, and Wood. *Journal of the American Statistical Association, 94*, 794-797.

Coull, B. A., Ruppert, D., & Wand, M. P. (2001). Simple incorporation of interactions into additive models. *Biometrics, 57*, 539-545.

Coull, B.A., Schwartz, J., & Wand, M. P. (2001). Respiratory health and air pollution: Additive mixed model analyses. *Biostatistics, 2*, 337-349.

Davidian, M., & Giltinan, D. M. (1995). *Nonlinear Models for Repeated Measurement Data.* Chapman & Hall, New York.

Diggle, P. J., & Hutchison, M. F. (1989). On spline smoothing with autocorrelated errors. *Australian Journal of Statistics, 31*, 166-168.

Diggle, P. J., Liang, K. Y., & Zeger, S. L. (1994). *Analysis of Longitudinal Data.* Oxford: Oxford University Press.

Eilers, P. H. C., & Marx, B. D. (1996). Flexible smoothing with B-splines and penalties (with discussion). *Statistical Sciences, 11*, 89-121.

Fan, J. & Zhang, J. T. (2000). Two-step estimation of functional linear models with applications to longitudinal data. *Journal of the Royal Statistical Society, Series B, 62*, 303-322.

Hoover, D. R., Rice, J. A., Wu, C. O., & Yang, L. P. (1998). Nonparametric smoothing estimates of timing-varying coefficient models with longitudinal data. *Biometrika, 85*, 809-822.

Laird, N. M. & Ware, J. H. (1982). Random effects models for longitudinal data. *Biometrics, 38*, 963-974.

Lederman, M. M., Connick, E., Landay, A, et al. (1998). Immunologic responses associated with 12 weeks of combination antiretroviral therapy consisting of Zidovudine, Lamivudine and Ri-tonavir: Results of AIDS clinical trials group protocol 315. *The Journal of Infectious Diseases, 178*, 70-79.

Liang, H., Wu, H. L., & Carroll, R. J. (2003). The relationship between virologic and immunologic responses in AIDS clinical research using mixed-effects varying-coefficient models with measurement error. *Biostatistics, 4*, 297-312.

Lin, X. & Carroll, R. J. (2000). Nonparametric function estimation for clustered data when the predictor is measured without/with error. *Journal of the American Statistical Association, 95*, 520-534.

Parise, H., Ruppert, D., Ryan, L., & Wand, W. P. (2001). Incorporation of historical controls using semiparametric mixed models. *Applied Statistics, 50*, 31-42

Pinheiro, J. C. & Bates, D. M. (2000). *Mixed-Effects Models in S and S-PLUS*, New York: Springer.

Rice, J. A. & Silverman B. W. (1991). Estimating the mean and covariance structure nonpara-metrically when the data are curves. *Journal of the Royal Statistical Society, Series B, 53*, 233-243.

Rice, J. A. & Wu, C. O. (2001). Nonparametric mixed effects models for unequally sampled noisy curves. *Biometrics, 57*, 253-259.

Ruppert, D. (2002). Selecting the number of knots for penalized splines. *Journal of Computational and Graphical Statistic (in press)*.

Ruppert, D. & Carroll, R. (1999). Spatially-adaptive penalties for spline fitting. *Australian and New Zealand Journal of Statistics, 42*, 205-253.

Shi, M., Weiss, R. E., & Taylor, J. M. G. (1996). An analysis of pediatrics CD4 counts for acquired immune deficiency syndrome using flexible random curves. *Applied Statistics, 45*, 151-163.

Vonesh, E. F. & Chinchilli, V. M. (1996), *Linear and nonlinear models for the analysis of repeated measurements*, New York: Marcel Dekker, Inc.

Wang, Y. D. (1998a). Smoothing spline models with correlated random errors. *Journal of the American Statistical Association, 93*, 341-348.

Wang, Y. D., (1998b). Mixed-effects smoothing spline ANOVA. *Journal of the Royal Statistical Society, Series B, 60*, 159-174.

Welsh, A., Lin, X., & Carroll, R. J. (2002). Marginal longitudinal nonparametric regression: Locality and efficiency of spline and kernel methods. *Journal of the American Statistical Association, 97*, 482-493.

Wu, C. O., Chiang, C. T., & Hoover, D. R. (1998). Asymptotic confidence regions for kernel smoothing of a varying-coefficient model with longitudinal data. *Journal of the American Statistical Association, 93*, 1388-1402.

Wu, H. & Zhang, J. T. (2002). Local polynomial mixed-effects models for longitudinal data. *Journal of the American Statistical Association, 97*, 883-897.

Zeger, S. L. & Diggle, P. J. (1994). Semiparametric models for longitudinal data with application to CD4 cell numbers in HIV seroconverters. *Biometrics, 50*, 689-699.